PROJECT – GROUP K

Giuliano Amadei, Tommaso Borelli, Maria Vittoria Di Cioccio, Riccardo Paolantoni.

Database Summary.

Our database contains data of almost every single athlete to have participated in the Olympics from its rebirth in 1896 until 2016. Due to the Covid-19 pandemic and the postponing of the 2020 Olympics to 2021, it is too recent to be included in the dataset.

The dataset originated from a table composed of 15 attributes and 257896 entries however, we decided to divide it into two tables named Seasons and Athletes.

Seasons contains four attributes called: *games, noc (Nation of choice), city and event*.

Athletes is composed of nine attributes: *name, noc, sex, age, height, weight, games, event* and *medal.*

Alongside our original dataset there was a complementary table called *noc\_regions*, which was composed of two attributes called *noc* and *region*. The attribute *region* defined every *noc* (e.g., CHN is defined as Chine in *noc\_reagions*).

We have chosen this dataset because it is very interesting due to multiple reasons;

1) It contains historical data.

* Due to it containing data revolving around political history like the independence of countries and political divides like e.g., the divide in Yemen between north and south.

2) It is a significant global event due to it being an event in which multiple different nations come together to compete in sport.

3) The dataset features events that are no longer practiced in the Olympics like art and sculpting.

## Workbench Instructions.

1. be sure to have created a MySQL connection.
2. Open MySQL workbench and create a new relation.
3. Right click on your relation and select “data import wizard” to import the tables in your schema: *seasons.csv, athletes.csv and noc\_region.csv.*
4. Repeat the previous steps for the three tables.

## Python Instructions.

1. When importing our *.sql* files, IT IS VERY IMPORTANT to import all files into the python working directory so as to avoid the need to specify the file path within the submitted python file;

If the files are not imported into the python working directory make sure to insert the file paths in the first string value after “open”.

1. Once the files are in the working directory no other steps need to be followed to run the file.
2. RUN ONCE READY.
3. If there is an exceedingly long run time with no changes in the console, interrupt the process and restart your computer.